Neural dynamics driving audio-visual integration in autism

Luca Ronconi[1,2*], Andrea Vitale[3*], Alessandra Federici[3,4], Noemi Mazzoni[3,5,6], Luca Battaglini[7,8], Massimo Molteni[9] & Luca Casartelli[3]

1. *School of Psychology, Vita-Salute San Raffaele University, Milan, Italy*
2. *Division of Neuroscience, IRCCS San Raffaele Scientific Institute, Milan, Italy*
3. *Theoretical and Cognitive Neuroscience Unit, Child Psychopathology Department, Scientific Institute IRCCS Eugenio Medea, Bosisio Parini, Italy*
4. *IMT School of Advanced Studies Lucca, Lucca, Italy*
5. *Laboratory for Autism and Neurodevelopmental Disorders, Center for Neuroscience and Cognitive Systems, Istituto Italiano di Tecnologia, Rovereto, TN, Italy*
6. *Department of Psychology and Cognitive Science, University of Trento, Rovereto, Italy*
7. *Department of General Psychology, University of Padova, Italy.*
8. *Department of Physics and Astronomy "Galileo Galilei", University of Padova, Italy.*
9. *Child Psychopathology Department, Scientific Institute IRCCS Eugenio Medea, Bosisio Parini, Italy*

*=*equally contributing authors*

Corresponding author:

Luca Ronconi, PhD.

Contact information: School of Psychology, Vita-Salute San Raffaele University, Via Olgettina 58, 20132, Milan (Italy). Email: ronconi.luca@unisr.it; Phone: +390226434887.

Abstract

Audiovisual (AV) integration plays a crucial role in supporting social functions and communication in autism spectrum disorder (ASD). However, behavioral findings remain mixed and, importantly, little is known about the underlying neurophysiological bases. Studies in neurotypical adults indicate that oscillatory brain activity in different frequencies subserves AV integration, pointing to a central role of: i) individual alpha frequency (IAF), which would determine the width of the cross-modal binding window; ii) pre-/peri-stimulus theta oscillations, which would reflect the expectation of AV co-occurrence; iii) post-stimulus oscillatory phase reset, which would temporally align the different unisensory signals. Here, we investigate the neural correlates of AV integration in children with ASD and typically developing (TD) peers, measuring electroencephalography during resting state and in an audiovisual integration paradigm. As for neurotypical adults, AV integration dynamics in TD children could be predicted by the IAF measured at rest and by a modulation of anticipatory theta oscillations at single-trial level. Conversely, in ASD participants AV integration/segregation was driven exclusively by the neural processing of the auditory stimulus and the consequent auditory-induced phase reset in visual regions, suggesting that a disproportionate elaboration of the auditory input could be the main factor characterizing atypical AV integration in autism.

Introduction

Multisensory processes represent a "building block" for higher-order functions essential for communication and social interaction (Spence & Squire, 2003; Murray et al. 2016). Recent evidence underlines the critical role of sensory peculiarities in autism spectrum disorder (ASD) symptomatology (Robertson and Baron-Cohen, 2017; Lai et al., 2020), with a special emphasis on multisensory processing anomalies (Wallace and Stevenson, 2014; Feldman et al. 2018; Zhou et al. 2018). Some studies report that, relative to their TD peers, children with ASD show a reduced susceptibility to audio-visual illusory phenomena such as the McGurk effect and the sound-induced double flash illusion (for review see Feldman et al. 2018; Zhou et al. 2018). Also, there is evidence that the "pip-pop effect", where an irrelevant sound can facilitate the detection of a target stimulus or features change, is reduced in ASD (Collignon et al., 2013). It has been also hypothesized that participants with ASD exhibit a wider extension of the temporal binding windows for audio-visual information of speech-like stimuli, and this seems to depend at least partly on their basic temporal acuity across vision and audition that can be measured even with simple stimuli like flashes and beeps (Stevenson, 2014; Feldman et al. 2018; Zhou et al. 2018). However, findings are mixed, with some reports of intact multisensory integration in ASD (Feldman et al. 2018). Several aspects such as age (Beker et al. 2018), type of task (Chan et al. 2016) and nature of the stimuli utilized in the tasks (e.g., speech vs. non-speech; see Stevenson et al. 2016) have been shown to influence multisensory integration in ASD, but which of them assume a preeminence in explaining multisensory integration peculiarities is not clear (Feldman et al. 2018; see also Zhou et al. 2018; Ainsworth et al. 2021). Notably, despite some behavioral evidence link ASD to multisensory integration/segregation unbalances, with potential important cascade effects in determining speech processing abilities, little is known about altered neurophysiological mechanisms behind these anomalies (Beker et al. 2018; Feldman et al. 2018; Zhou et al. 2018).

Any computation that involves communication between distinct brain areas (e.g. top-down/feedback or cross-sensory influences) leads to discrete or quasi-discrete processing windows, strongly

influencing the output of sensory computations (Pöppel, 2009). In addition, endogenous fluctuations of neural excitability modulate the processing of incoming sensory stimuli and influence perceptual, motor, and cognitive processes in different ways (Morillon and Schroeder, 2015; VanRullen, 2016a). In multisensory processing, balancing temporal integration and segregation has a crucial impact on how sensory signals from different modalities are organized in our perceptual experience. This integration/segregation balance has specific neural signatures tightly linked to different brain oscillatory patterns (Bauer et al. 2020; Keil and Senkowski, 2018). Primarily, temporal cross-modal integration is dependent on endogenous or pre-stimulus neural rhythms: individuals with a higher alpha peak frequency in the resting-state or in the pre-stimulus period exhibit better temporal segregation across sensory modalities (Cecere et al., 2015; Keil and Senkowski, 2017; Cooke et al., 2019; Venksus and Hughes, 2021). In addition, multisensory integration is supposed to involve cross-modal influence between cortical areas, with changes of pre-stimulus low/mid frequencies activity reflecting prediction of the timing of auditory stimuli based on visual dynamics (Keil and Senkowski, 2018; Ten Oever and Sack, 2015). Moreover, increasing evidence in neurotypical adults suggests that auditory stimuli can "reset" the phase of oscillatory activity in the visual cortex (Lakatos et al., 2009; Romei et al., 2012; Mercier et al., 2013), which would be important for aligning rhythmic brain activity of auditory and visual areas and for opening a temporal window where audio-visual integration can occur. For example, auditory stimulation can influence the processing of a concurrent visual stimulus, leading to an increased inter-trial phase coherence (ITC) in the visual cortex in theta (4–8 Hz), alpha (8–12 Hz), and beta (13-20 Hz) bands (Mercier et al., 2015; Romei et al., 2012, Thézé et al., 2020).

The present study employs electroencephalography (EEG) to investigate whether differences in endogenous, pre-stimulus, or stimulus-evoked oscillatory activity could underlie atypical multisensory integration/segregation balance in individuals with ASD. As a proxy for testing multisensory dynamics in ASD, we used a largely employed audio-visual integration paradigm, i.e. the stream-bounce illusion (Sekuler et al., 1997), where two moving objects approached each other,

overlapped, and diverged again following a constant trajectory. A sound-dependent change in visual motion perception is the key aspect for testing audiovisual (AV) integration vs. segregation. Indeed, if a bounce-like sound occurs around the time of objects overlap, the moving objects can be perceived either as passing each other ('stream' trials, indexing segregation of the two modalities) or bouncing off ('bounce' trials, indexing integration). In this bistable scenario, the divergent perceptual responses to the same physical stimulation can be used to study audio-visual integration/segregation dynamics and their putative anomalies in ASD. In particular, we explored the hypothesis that one or more neural signatures described in the neurotypical population could be altered or unbalanced in children with ASD, such as: i) a weaker or absent link between spontaneous alpha rhythms and AV integration window, which would suggest a poor reliance on endogenous oscillatory activity and/or a more pronounced stimulus-driven multisensory sampling; ii) an absence of pre-stimulus power/phase differences between AV integration vs. segregation in low-/mid- frequency oscillations that would index a marginal reliance on anticipatory signals of AV co-occurrence; iii) a different impact of the sound and of the sound-induced phase reset, which would point to an atypical balancing and temporal alignment of the different unisensory signals.

Methods

Participants

Twenty-two children with ASD (4 females) and 21 TD (3 females) children took part in the present study. Children with ASD were recruited from the Scientific Institute IRCCS "E. Medea" (Bosisio Parini, Italy) according to the following criteria: full-scale IQ (FSIQ) > 70 as measured by the WISC III or WISC IV (Wechsler, 1993); absence of gross behavioral problem; normal/corrected-to-normal vision and hearing; absence of drug therapy; absence of diagnosis of attention deficit or hyperactivity disorder (American Psychiatric Association, 2013). Diagnosis of ASD was made by licensed clinicians in adherence to the DSM-IV or the DSM V criteria and confirmed with the Autism

Diagnostic Observation Scale (ADOS; Lord et al., 2002). The TD children were recruited from local schools in the same geographical area, and were administered with Block design and Vocabulary subtests of the WISC-IV battery. Those measures were taken as indexes for nonverbal and verbal abilities and were used as matching criteria with the ASD group. The two groups were matched for chronological age ($t_{(31.80)}$ = -1.52, p = 0.137) and for Block design performance ($t_{(29.56)}$ = 1.34, p = 0.189), but they differed for Vocabulary performance ($t_{(31.27)}$ = 3.52, p = .001). Since the task that we used in the present experiment consisted in a simple audio-visual integration task with no verbal material, we are inclined not to consider the difference in the Vocabulary task as a bias for the obtained results. Finally, parents of all children who took part in the present study were asked to fill in the Social Communication Questionnaire (SCQ). Participants with ASD, as expected, scored significantly higher than TD children both in "Current" ($t_{(20.10)}$ = - 4.63, p < .001) and "Life time" forms ($t_{(15.67)}$ = -7.56, p < .001), meaning that they showed a higher presence of symptoms in the social and communicative domains.

In order to obtain a balanced number of stream and bounce trials for the EEG analyses, we defined inclusion criteria for participants based on their illusory rate (i.e. percentage of bounce reported over the total number of trials), setting a higher cutoff at 85% and a lower cutoff at 15%. According to these criteria, two children with ASD and two TD children presented an unbalanced perceptual bistability with a percentage of bounce responses in the synchrony condition higher than 85% and were therefore excluded from further analysis. Moreover, three children from the ASD group and two children from the TD group were excluded from the analysis because their EEG data were excessively contaminated by artifacts (more than 35% of epochs removed). According to these criteria, a total of 17 children with ASD and 17 children with TD were considered in the final behavioral and EEG analysis.

The study was approved by the ethical committee of the Scientific Institute IRCCS "E. Medea". Informed consent was obtained from every child and his/her parents, in accordance with the declaration of Helsinki.

144 [Insert Table 1 about here]

	ASD (N=17; 4F)	TD (N=17; 2F)	p-value
	mean (SD)	mean (SD)	
Age	12.47 (1.93)	11.5 (1.78)	0.137
Block Design	11.59 (3.18)	12.88 (2.37)	0.189
Vocabulary	9.06 (2.11)	11.82 (2.45)	0.001
Social Communication Questionnaire (SCQ) – Current	12.60 (5.97)	4.69 (2.94)	<.001
Social Communication Questionnaire (SCQ) - Lifetime	21.64 (8.31)	4.00 (2.85)	<.001
ADOS Classification	ADOS 2: Autism Spectrum (N=3) Autism (N=1)	--	--
	ADOS: Autism Spectrum (N=10) Autism (N=3)		

Table 1. *Descriptive statistics of participants (ASD = autism spectrum disorder; TD = typically developing; SCQ = Social Communication Questionnaire; ADOS = Autism Diagnostic Observation Scale).*

145

146 *Apparatus, stimuli and procedure*

147 The visual stimuli were presented on a Philips 19SL monitor (60 Hz, resolution 1280x960) and
148 consisted of two white dots with a diameter of 1.3 deg, moving horizontally on a grey background at
149 a velocity of 0.13 deg per frame. One dot moved from the left to the right side of the screen, the other
150 dot moved from the right to the left side of the screen. The dots appeared on the horizontal alignment
151 and their movements started simultaneously at an eccentricity of 11.7 deg and disappeared at a
152 position of 7.8 deg of eccentricity in the opposite hemifield. The two dots overlapped each other at
153 the center of the screen after 1500 ms from the stimulus onset. In total, the duration of the movement
154 lasted for 2500 ms.

155 The auditory stimulus was characterized by an abrupt amplitude attack and a gradual decay with an
156 intensity peak occurring within ~10 msec after the sound onset and a duration of 500ms. The bounce-
157 like sound was presented at a sampling rate of 44100 Hz through two speakers (Hercules XP 2.0 60

158 DJ SET) located laterally to the monitor where the visual stimuli were displayed. The sound could be
159 delivered: 150 ms before the complete overlapping of the two dots ('pre-coincidence' condition); 67
160 ms before the complete overlapping of the two dots corresponding to the moment of the dots approach
161 ('synchronous' condition); or 184 ms after the complete overlapping of the two dots ('post-
162 coincidence' condition) (see Figure 1A).
163 The latencies of the sound in the three conditions were selected in order to elicit perceptual bistability
164 based on previous pilot observations. In particular, the timing of the events in the synchronous
165 condition, with the sound appearing 67 ms before complete overlapping, was expected to elicit a
166 balanced number (i.e. ~50%) of stream/bounce percepts. Participants sat in a dark room at a distance
167 of 50 cm from the monitor. An experimenter was next to the participants for the entire duration of the
168 experiment to collect the children's responses and to reinforce their efforts or to support the children
169 with ASD when needed. In order to minimize the eye movement artifacts, the participants were
170 invited to blink only during the inter-trial interval. The entire session was divided in six blocks of 60
171 trials each, for a total of 120 trials per condition. The three conditions appeared with equal probability
172 and were presented in a random order. At the end of every trial, participants were asked to report
173 verbally whether they saw the two dots "streaming" or "bouncing" each other, and the experimenter
174 pressed the corresponding button on the computer keyboard (to minimize motor artifacts). The visual
175 stimuli were displayed and responses were recorded using the MATLAB Psychtoolbox (Brainard,
176 1997; Pelli, 1997).
177
178
179 [Insert Figure 1 about here]

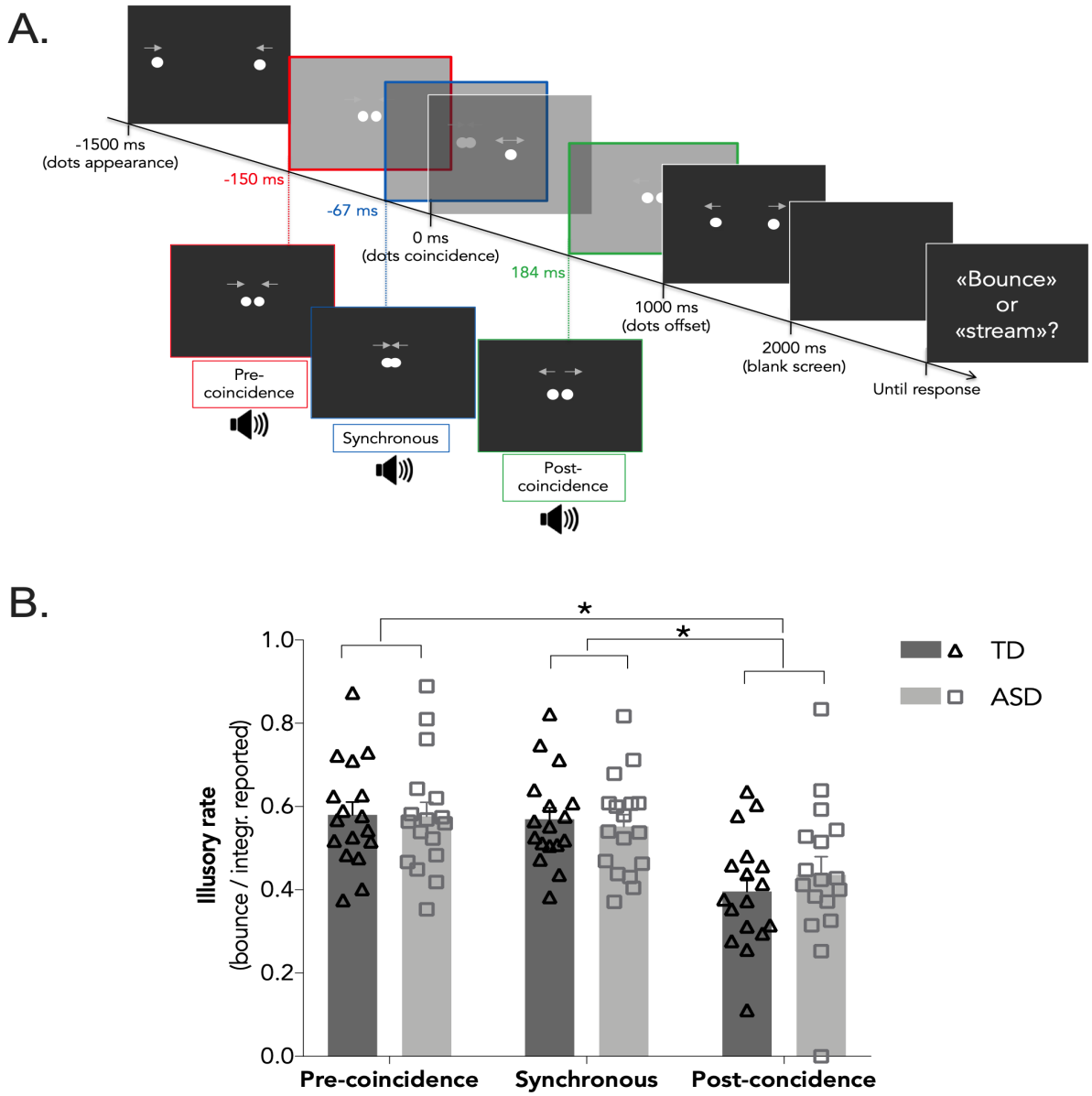


Figure 1. Representation of the experimental paradigm and behavioral results. (A) Example of an audiovisual bounce-inducing trial employed in the study. Two dots appear simultaneously at both sides of the screen and start to move horizontally until they overlap each other in the center of the screen (1500 ms after the stimulus onset). According to the three trial conditions, labelled as 'pre-coincidence', 'synchronous' or 'post-coincidence' conditions, a bounce-like sound is delivered respectively 150 ms before, 67 ms before or 184 ms after the complete overlap of the two dots. At the end of each trial, participants were asked to verbally report whether they perceive the two dots as "streaming" or "bouncing" each other. (B) The illusory rate reported in the bar graph for each group (TD and ASD) and conditions was computed as the rate of bounce trials perceived over the total number of trials. The occurrence of the auditory event before or in synch with the dots overlap (pre-coincidence and synchronous conditions) induced an illusory bounce perception in more than half of the trials, while in the post-coincidence condition perception was biased toward the stream outcome. Triangles and squares represent individual values.

EEG recording and pre-processing

Neural activity of children was recorded by using the Electrical Geodesics system with a 64-channel Hydrocel Geodesic Sensor Net (Electrical Geodesics, Inc.). The EEG system collected the data with

196 a sampling rate of 500 Hz and automatically performed an analog filter between 0.01 and 100Hz
197 (before digitization). Offline, data were band-pass filtered between 0.05 and 80 Hz (non-causal IIR,
198 order = 2) and a notch-filter at 50 Hz (non-causal Parks-McClellan Notch, order = 180) was applied.
199 Then data were downsampled to 250 Hz and the reference was recomputed to the average. The pre-
200 processing steps, including filtering, artifact correction/removal and epoching, were performed in
201 MATLAB using EEGLAB (Delorme and Makeig, 2004) and ERPLAB (Lopez-Calderon and Luck,
202 2014).
203
204 *Behavioural analysis*
205 For each participant, the percentage of bounce responses over the total amount of responses (i.e.
206 illusory rate) was calculated for each condition (i.e. pre-coincidence, synchronous, post-coincidence)
207 and used as the dependent variable. Such illusory rate was analyzed with a $2x3$ mixed ANOVA with
208 Group (ASD and TD) as between-subject factor and Condition (pre-coincidence, synchronous, post-
209 coincidence) as within-subject factor. All behavioral analyses were performed with the software R
210 (version 3.4.4).
211
212 *EEG analysis - Event-related potential (ERPs)*
213 EEG data were segmented in 3 sec epochs, 1.5 sec before and 1.5 sec after the dots complete overlap
214 (120 epochs for each of the three conditions). Spherical interpolation was performed on individual
215 bad channels when required (average number of interpolated channels: 2.21 in TD children; 3.38 in
216 children with ASD). Blinks and eye movements artifacts were identified and removed using an
217 independent component analysis (ICA). Furthermore, epochs containing voltage deviation exceeding
218 ±150 μV were removed. Following this procedure, on average 6.11% of epochs in TD children and
219 14.4% in children with ASD were removed.
220 The activity between -1500 ms and -1400 ms before the dots overlap was considered as baseline for
221 ERPs analyses. Since the task consisted in an audio-visual integration, analyses were focused over an

occipital cluster (electrodes: E33, E34, E36, E38) and two clusters of temporal electrodes: a right temporal cluster (electrodes: E46, E48, E49, E52) and a left temporal cluster (electrodes: E22, E24, E25, E26). The activities of these electrodes were averaged before the application of statistical analyses.

To detect differences between the two groups when they perceived the dots bouncing or streaming, we tested the presence of a possible $2x2$ interaction by using nonparametric cluster-based permutation tests (N. of permutations=10000); these interactions were tested separately in each experimental condition (pre-coincidence, synchronous, post-coincidence) in all time points ranging from -800 to +800 relative to the dots overlap. This method is data-driven and allows to test point-by-point the significant interaction between Group (ASD vs TD) and the Type of percept (bounce vs stream) in the entire time interval and within all 3 clusters of interest while controlling for multiple comparisons with the cluster-based correction method (Maris and Oostenveld, 2007). Specifically, to test such cluster-based interaction we computed the within-group bounce minus stream amplitude difference and used cluster-based permutation tests to compare the two groups.

Furthermore, if there were time windows that resulted significant from the cluster-based permutation tests, we further explored the existence of an interaction using mixed design parametric ANOVAs with the ERP mean amplitude as dependent variable, Group (ASD vs TD) as between-group factor, and Type of percept (bounce vs stream) as within-group factor.

EEG analysis - Time-frequency decomposition and oscillatory power analysis

Single trial EEG data were time-frequency transformed with a wavelet analysis as implemented in the *newtimef.m* function from the EEGLAB toolbox. Output frequencies ranged from 3 to 80 Hz in 100 logarithmically spaced steps, while the wavelet width scaled with frequency values from 4 cycles at 3 Hz to 25 cycles at 80 Hz to obtain a good trade-off between frequency and time resolution (particularly important at lower frequencies when inspecting any pre-stimulus effect). As baseline we

used the interval from -1000 to -700 ms, right after the neural response evoked by the visual stimulus onset.

For (second level) statistical evaluation, similarly to what has been described for the ERPs analyses and in order to specifically investigate the oscillatory activity related to the AV integration, in each time-frequency point, nonparametric $2x2$ cluster-based permutation tests were performed to test the presence of significant interaction between Group (ASD vs TD) and Type of percept (bounce vs stream). To this aim we computed the bounce minus stream power difference and used cluster-based permutation tests to compare the two groups. In case of significant frequency x time clusters, we further applied a mixed design $2x2$ parametric ANOVA with factors Group (TD vs. ASD) and Type of percept (bounce vs. stream).

EEG analysis - Inter-trial Phase Coherence analysis

The extent of phase concentration across trials was quantified on the bases of the inter-trial phase coherence (ITC) (Tallon-Baudry et al., 1996). This measure is computed with the complex output of the wavelet transform and provides an indirect representation of the systematic increase in phase concentration across trials, with possible values ranging from 0 (complete absence of phase coherence) to 1 (perfect phase locking). As for power analyses described above, we first investigated a $2x2$ interaction between Group and Type of percept by means of permutation tests, computing the bounce minus stream ITC difference and using cluster-based permutation tests to compare these measures in the two groups. In case of significant frequency x time clusters, the Group x Type of percept interaction was further explored with a parametric ANOVA.

In particular, we were interested in detecting sustained levels of phase coherence in the post-stimulus time window without a concomitant increase of power, since such effect would potentially indicate a phase resetting mechanism (Mercier et al., 2013; Makeig et al., 2002; Kambe et al., 2015).

EEG analysis - Phase Opposition Indexes

A final step of analysis was devoted to quantify the relationship between the perceptual outcomes and the ongoing oscillatory phase in the pre-stimulus interval. Since the bistable paradigm employed consisted in the presentation of exactly the same stimuli, the divergent perceptual outcome (i.e. bounce or stream) might be determined by the instantaneous phase at which such cognitive process is engaged.

A variety of analytical methods have been adopted to measure the impact of the neural oscillatory phase on the trial-by-trial fluctuations, and all of them are based on the assumption that the phase coherence across a subset of trials (in a specific condition) should exceed the ITC sum for all the trials (Van Rullen, 2016b). Among these procedures, the "Phase Opposition Product" (POP) has been independently employed in a large number of studies (Han and VanRullen, 2017), and its original formulation was described as:

$$POP = (ITC_{bounce} \cdot ITC_{stream}) - ITC_{all}\text{^}2$$

where ITC_{all} was calculated considering all trials independently from their perceptual outcome.

The values of the POP index are bounded theoretically between 0 and 1; values approximating 1 indicate that the ITC of both trial conditions exceeds the overall ITC, as for example, when the phase of the two subsets of trials (bounce and stream) are consistently clustered around two different angles. On the contrary, a POP value close to zero may occur only when the phase coherence for both trial conditions is low, suggesting a random/uniform phase distribution.

Between groups, the statistical significance of POP measure was assessed through a permutation test with a cluster correction for multiple comparison in the pre-stimulus time window ([-700, 0 ms]) and in a frequency range covering theta, alpha, and low beta band (4-20 Hz) where phase information has been associated to multisensory processing (Ten Oever and Sack, 2015; Keil and Senkowski, 2017). To better interpret the directionality of the POP within each group, we conducted a point-by-point Rayleigh test, separately on bounce and stream trials, to identify time windows with a significant phase concentration (i.e. not distributed uniformly around the circle). Moreover, a circular analogue

of the two-samples t-test was used to assess whether the phase angle between the two perceptual outcomes was comparable or not (Berens, 2009).

Individual alpha peak at rest and its correlation with the stream-bounce illusion

Resting-state EEG data were obtained from 3 minutes of eyes-closed recording performed before the stream-bounce task. Data underwent the same pre-processing steps described above, with the exception that the band-pass filter was applied between 0.05 and 30 Hz. Data were visually inspected to remove data segments contaminated by muscular or ocular artifacts before extracting the Fast Fourier Transform (FFT) spectrum. The individual alpha frequency (IAF) peak was calculated as the frequency corresponding to the maximum FFT power values between 8 and 12 Hz, considering the average of all channels in a cluster (see "Event-related potential (ERPs)" section). IAF values calculated in this way were correlated both with measures of ASD symptomatology - as indexed by the SCQ scores - and with the illusory rate in the stream-bounce task (i.e. the rate of trials perceived as bounce over the total number of trials). We used partial correlations to account for the variability in participants' age.

To further assess that the correlation results were not driven by some group differences in the distribution of the individual alpha peak or by a less pronounced alpha peak as compared to the 1/f background slope, we compared the IAF extracted for ASD and TD groups.

Results

Behavioral performance

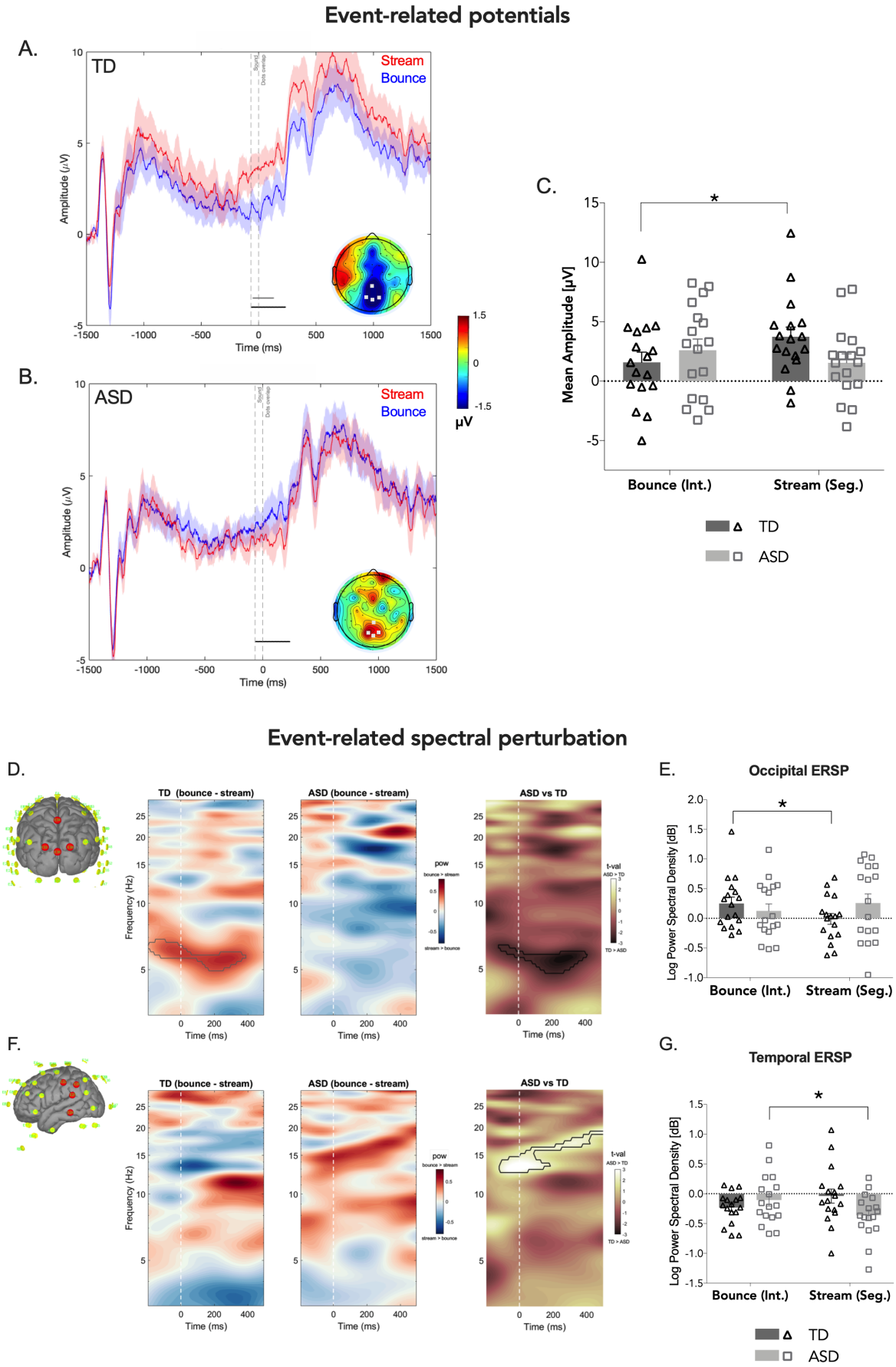
The assumptions for the ANOVA were met, as shown by the Shapiro test (all participants: $W = 0.984$, $p = 0.240$; TD: $W = 0.993$, $p = 0.986$; ASD: $W = 0.966$, $p = 0.156$) and by the Levene's Test for Homogeneity of Variance between Groups and Conditions that was not significant ($F_{(5,96)}=0.329$, $p = 0.894$). Results of the ANOVA showed a significant main effect of the factor Conditions ($F_{(2,64)}=20.07$, $p < .001$, $\eta^2_G=0.223$; Figure 1B), with a lower illusory rate in the post-coincidence condition compared to the synchronous ($p < .001$) and pre-coincidence ($p < .001$) conditions. The occurrence of the auditory event slightly before or in synchrony with dots collision induced an illusory percept (i.e. bounce perceived) in more than half of the trials, while in the post-coincidence an opposite pattern emerged, with illusory reports that were below 50% (Figure 1B). Neither the effect of Group (ASD vs. TD) nor the interaction between Group and Conditions resulted significant.

ERPs results

Results of permutation tests revealed a significant cluster-corrected $2x2$ interaction (Group x Type of percept) in the synchronous condition in the occipital channels, in a peri-stimulus time window ranging from -60 ms to 232 ms (cluster-corrected p-value= .0096; Figure 2A-B); when further explored we found that the direct comparisons between bounce and stream trials was significant only in TD children from -48 ms to 128 ms (cluster-corrected p-value= .021; Figure 2A-B). No significant interactions were found in the other cluster of channels nor in the other experimental conditions (pre-/post- coincidence) tested (see Supplementary Figure 1).

To further explore the significant interaction obtained in the synchronous condition, we performed a parametric ANOVA on the mean amplitude in this time interval (Figure 2C). The assumptions for the ANOVA were met, as shown by the Shapiro-Wilk normality test (W= 0.985, p= 0.615) and Levene's Test for Homogeneity of Variance ($F_{(3,64)}$= 0.763, p= 0.519). This ANOVA confirmed the result obtained with permutation tests by showing a significant interaction between Group and Type of percept ($F_{(1,32)}$= 11.104 , p= .002, η^2_G = 0.053). In particular, post-hoc comparisons revealed that only in the TD group ($t_{(32)}$= -3.385, p= .004; Figure 2C) the mean amplitude was significantly smaller in trials with AV integration (bounce perceived) as compared to trials where AV components were segregated (stream perceived). In the ASD group, the difference between stream and bounce mean amplitude was not significant ($t_{(32)}$= 1.453, p= 0.166).

[Insert Figure 2 about here]


Event-related potentials
A.
TD
Stream
Bounce
Amplitude (μV)
Time (ms)
B.
ASD
Stream
Bounce
Amplitude (μV)
Time (ms)
μV
C.
Mean Amplitude [μV]
Bounce (Int.)
Stream (Seg.)
TD
ASD
Event-related spectral perturbation
D.
TD (bounce - stream)
ASD (bounce - stream)
ASD vs TD
Frequency (Hz)
Time (ms)
pow
t-val
E.
Occipital ERSP
Log Power Spectral Density [dB]
Bounce (Int.)
Stream (Seg.)
F.
TD (bounce - stream)
ASD (bounce - stream)
ASD vs TD
G.
Temporal ERSP
Log Power Spectral Density [dB]
Bounce (Int.)
Stream (Seg.)
TD
ASD

350

Figure 2. Event-related potentials (ERPs) and event-related spectral power in bounce vs. stream trials. (A, B) Neural evoked response (ERPs) in occipital channels (E33, E34, E36, E38) and related topographical scalp maps (bounce-stream difference) for the synchronous condition in the TD and ASD group. Trials were split according to the perceptual outcome reported, and a 2*x*2 interaction between group (ASD vs. TD) and Type of percept (bounce vs. stream) emerged in a time window around the two dots overlap (black horizontal lines displayed in both panel A and B). In TD children, a significant difference between bounce (integration) and stream (segregation) trials was detected in the occipital channels within this time window (from -48 ms to 128 ms; gray horizontal line), while no significant differences emerged in the ASD group. (C) Mean amplitude of the peri-stimulus interval as a function of the group and Type of percept: in TD group the neural response was smaller in AV integration trials (bounce perceived) as compared to AV segregation trials (stream perceived). Such modulation was absent in the ASD group. Triangles and squares represent individual values. (D) Time-frequency plot displaying for each group (TD and ASD) the difference in oscillatory amplitude (i.e. power) between stream trials (considered as the control condition) and the bounce trials, as computed over the occipital cluster of channels. On the right side a t-value map shows the significant cluster-corrected interaction effect in the theta band (~ 4-8 Hz) in a time window ranging from ~ -100 to 400 ms relative to the dots overlap. (E) Bar plot showing occipital power (ERSP) in the time-frequency window where a significant interaction emerged as a function of group and perception of AV integration (bounce) or segregation (stream) (*=p<.05; triangles and squares represent individual values). (F) Same analysis as in (D) performed on the electrodes located on the left temporal area revealed a significant difference between groups in the alpha/low-beta frequency range (12-20 Hz) from ~ -100 to 500 ms relative to the dots overlap. (G) Bar plot showing left temporal power (ERSP) in the time-frequency window where a significant interaction emerged as a function of group and perception of AV integration (bounce) or segregation (stream) (*=p<.05; triangles and squares represent individual values).

Oscillatory power analysis

In the synchronous condition, the cluster-based permutation analysis conducted on the spectral power difference revealed two distinct interactions between Group x Type of percept in two different clusters of channels, but in a comparable time window.

The first interaction in the synchronous condition emerged in the occipital cluster in the theta band (~ 4-8 Hz, ~[-100, 400 ms]; Figure 2D). This interaction was further studied with an ANOVA that confirmed a significant interaction effect ($F_{(1,32)}$=5.621, p= .024; Figure 2E). Planned comparisons in this time-frequency window showed a prominent increase in the theta power in bounce trials (integration) as compared to stream (segregation) trials in the TD group ($t_{(16)}$= 2.44, p= .027; Figure 2E). Such modulation was not significant in participants with ASD.

The second interaction in the synchronous condition was found in the left temporal cluster in the alpha and low-beta band (12-20 Hz, ~[-100, 500 ms]; Figure 2F). Also in this case it was further confirmed with the ANOVA ($F_{(1,32)}$=6.652, p= .015; Figure 2G). Planned comparisons in this time x frequency window showed a pronounced increase of the oscillatory power in the illusory trials in the

ASD group ($t_{(16)}$= 2.362, p= .031; Figure 2G). This power difference between the two experimental conditions was not observed in the TD group.

It is important to note that in the investigated time window, the time points where significant results emerged might be affected by a temporal contamination (or "leakage") induced by the Wavelet analysis. For this reason, it cannot be excluded that the significant time x frequency region close to the event onset was affected by some post-stimulus activity, therefore we decided to describe this result as a "peri"-stimulus effect.

Finally, there were no significant $2x2$ interactions, tested at each time-frequency point with the nonparametric permutation tests and cluster-based correction nor in the pre-coincidence condition nor in the post-coincidence condition in any of the channels group considered.

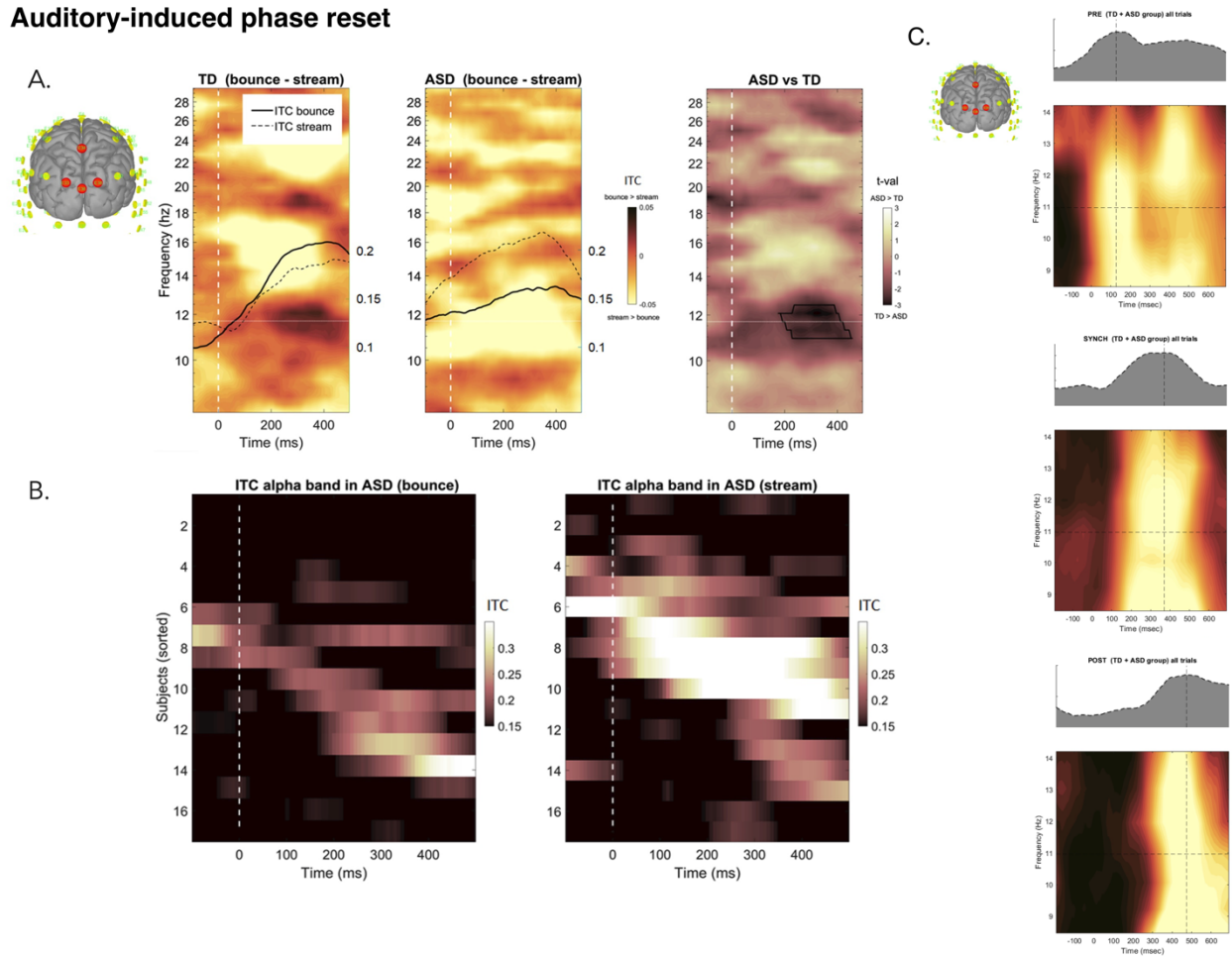
Inter-trial phase coherence (ITC)

The oscillatory activity between -100 and 500 ms time window was investigated also in terms of ITC. At the descriptive level, both groups presented as expected an increase of ITC after the auditory event. Importantly, the $2x2$ interaction tested with permutation tests in the synchronous condition was strictly speaking not significant, but there was a trend in the alpha/low-beta frequency range (11-13 Hz, ~[200, 400 ms]) for the occipital cluster (cluster-corrected p-value= .0604; Figure 3A). When further studied with a Group x Type of percept parametric ANOVA, the interaction effect in the occipital cluster resulted significant ($F_{(1,32)}$= 5.302, p= .028). The interaction was driven by a significant enhancement of ITC in stream as compared to bounce trials in the ASD group ($t_{(16)}$= -3.291, p= .005; Figure 3A; see also the distribution of the individual coherence peaks illustrated in Figure 3B). On the contrary, in the TD group, the ITC did not differ significantly between conditions (integration/bounce and segregation/stream), although the direction was comparable with previous literature, i.e. ITC in bounce trials was higher than in stream trials.

In the pre-coincidence condition, there were no significant $2x2$ interactions as revealed by permutation tests-based ANOVA.

In the post-coincidence condition, similarly to the synchronous condition, there was in the occipital channels a condition by group interaction which was only close to significance (cluster-corrected p= .067; Supplementary Figure 2), in a time-frequency range similar to that observed in the synchronous condition (16-22 Hz, ~[200, 400 ms]). When further studied with a Group x Type of percept parametric ANOVA, the interaction effect in the occipital cluster resulted significant ($F_{(1,31)}$= 8.137, p= .006); such interaction was driven by a significant difference between bounce and stream trials within the TD group, who presented an ITC increase in the bounce (integration) trials as compared to the stream trials (segregation) ($t_{(16)}$= 2.7, p= .043). This difference within the TD group is in line not only with what we expected from previous literature in typical adults, but also with the direction of the effect observed in the synchronous condition (i.e. ITC in bounce trials higher than in stream trials), albeit in that case it was not significant.

426 The increment in ITC observed in the alpha/low-beta band in the post-stimulus period, and
427 specifically its ramping curve, was observed in the absence of a concomitant power increase in the
428 same frequency range in both groups, that would constitute a major confound for the interpretation
429 of ITC results. Importantly, this is compatible with a key role for the auditory-driven resetting of the
430 ongoing phase as a neural coding mechanism (Mercier et al., 2013).
431 To further confirm that changes in the ITC of occipital areas in the present paradigm are driven by
432 the auditory input, we extracted the ITC for both groups considering all trials (without distinguishing
433 the perceptual outcome) as a function of the sound occurrence (pre-coincidence, synchronous and
434 post-coincidence). The rationale for this additional analysis is that if such phase synchronization in
435 occipital areas would be influenced by some salient auditory event, the latency of the ITC peak would
436 be strictly modulated by the temporal onset of the auditory event itself. This pattern was precisely
437 what we found and reported in Figure 3C.
438 A similar dependency between the auditory stimulus timing and ITC was not observed in the temporal
439 clusters (of both left and right hemisphere), suggesting that the observed phase reset originated in the
440 auditory areas and exerted its cross-modal effect in visual areas.
441
442 [Insert Figure 3 about here]



443 **Figure 3**. (A) (B) (C) Inter-trial Phase Coherence analysis in the occipital channels. (A) As in the previous figure,
444 since the primary purpose was to investigate the interaction effect between Group and Type of percept, the t-value
445 map on the right side shows the result of the permutation statistics computed on the bounce minus stream ITC
446 difference between groups. The two plots on the left displayed, for each group separately, the inter-trial phase

coherence when subtracting the stream (segregation) trials from the bounce (integration) trials. In particular, the two lines superimposed on the time-frequency map represented the course of the phase synchrony (ITC) after the auditory stimulation for a specific frequency in the alpha band (at 11.6 Hz): while in TD there was an increment in ITC that did not differ between conditions, in the ASD group the stream/segregation trials (dotted line) presented an increased ITC in the post-sound period as compared to bounce/integration trials. (B) Distribution of the individual coherence peaks in the ASD group in the alpha range identified through the cluster analysis for the bounce (left plot) and stream (right plot) conditions. (C) Inter-trial Phase Coherence comparison among trials in the pre-coincidence, synchronous and post-coincidence conditions. The time-coherence maps for the occipital channels illustrate the latency of the phase synchronization in the alpha band as a function of the auditory stimulus onset. In particular, the top panels displayed the time-course of the ITC at 11 Hz, with the ITC peak clearly modulated by the timing of the auditory stimulus. For this analysis the ITC values of both groups and all trials, independently from the perceptual outcome, were considered. (D) (E) (F) Phase Opposition Product (POP) between stream and bounce trials in the left temporal cluster during the pre-stimulus interval. (D) Permutation statistics pointed out a region of statistical significance of POP between TD and ASD groups in the theta (4-8 Hz) frequency range. (E) Anticipatory phase distribution for bounce (integration) trials, in black, and stream (segregation) trials, in red, for a specific frequency in the theta range (4.6 Hz) for the TD group. To further explore the directionality of the effect, the lower panel reported the extended pre-stimulus window from about -500 to -200 ms where the phase distribution for the bounce trials was not uniformly distributed (statistical threshold of 0.05 uncorrected for the Rayleigh test), differently from the random phase distribution of the stream trials. (F) In ASD participants such discrepancy between the phase distribution of the two types of percepts was absent: i.e. no timepoints in the pre-stimulus period had a significant different phase concentration for the bounce or the stream trials.

Phase opposition indices

In the synchronous condition, the permutation statistics highlighted for the left temporal cluster a significant difference between the two groups (~4-8 Hz, ~[-600, -200 ms]; cluster-corrected p-value= .0199; Figure 3D) in the POP localized in the pre-stimulus interval across the theta frequency range. When explored in detail the anticipatory oscillatory phase has an impact in determining the perceptual outcome only in the TD group. In particular, for the bounce trials the Rayleigh test revealed an extended time window (~[-500, -200 ms]) where the phase distribution was consistently concentrated around a common angle, contrarily to the uniform phase distribution in the stream trials. Such discrepancy in the phase concentration between the two conditions was furtherly assessed by the circular analogue of the two samples t-test (as shown in Figure 3E). Interestingly, in the ASD group an equivalent departure from uniformity for bounce and stream trials was absent, and no significant phase distribution difference between the two conditions was observed (Figure 3F).

Similar results were observed also in the pre-coincidence condition. Indeed, we found a significant POP difference in the temporal left channels between ASD and TD children in a similar time-

frequency range (7-10 Hz; (~[-450, -200 ms]; cluster p-value= .039; Supplementary Figure 3) and in the same direction, i.e. the POP of the TD group was higher than the POP of the ASD group.

We did not find any significant POP difference in the post-coincidence condition; in this case, however, we should point out that the distribution of stream/bounce responses (and the relative/corresponding number of trials) is farther away from the ideal bistability. Considering the need to equalize the number of trials entering the POP analysis to the condition with less trials, such unbalance trial distribution could undermine the sensitivity of this index.

Individual Alpha Peak (IAF) correlations with behavioural results and ASD symptomatology

The two groups did not show significant differences in the IAF peak in all three clusters of channels (all ps >.173). Importantly, according to the previous literature in TD adults, we found that in our group of TD children the illusory rate in the synchronous condition (i.e. percentage of bounce perceived) showed a significant negative correlation with the IAF, which was stronger in the occipital cluster (Occipital: $r_{(12)}$=-0.846, p<.001, see Figure 4A; Temporal left: $r_{(12)}$=-0.56, p=.037; Temporal right: $r_{(12)}$=-0.588, p=.027). These results suggest that TD participants with the slower alpha rhythm were the ones showing the stronger illusory effect, namely a temporal widening of the AV binding window resulting in an increased number of illusory percepts. Contrarily, in the ASD group these correlations between illusory rate and IAF were not significant (all ps> .161).

Results in the pre-coincidence condition showed an equivalent pattern. Indeed, in the TD group we observed a significant negative correlation between the IAF and the illusory rate, which was stronger in the occipital cluster (Occipital: $r_{(12)}$=-0.865, p<.001; Temporal left cluster: $r_{(12)}$=-0.57, p=.026; Temporal right cluster: $r_{(12)}$=-0.643, p=.01). Contrarily, in the ASD group these correlations between illusory rate and IAF were again not significant (all ps> .136). In the post-coincidence condition, no significant correlation emerged neither in the TD (all ps>.189) nor in the ASD group (all ps> .136).

Finally, when considering the possible relationship between ASD symptomatology as measured by the SCQ (Current version) and IAF, we found significant negative correlations in the ASD for the two clusters of temporal channels (Temporal left: $r_{(9)}$=-0.626, p=.039; Temporal right: $r_{(9)}$=-0.72, p=.013; Figure 4B), showing that more severe social and communicative difficulties were associated with lower temporal IAF in ASD participants. The same correlations resulted significant also when considering all participants independently from the group (Temporal left IAF: $r_{(24)}$= -.477, p=.014; Temporal right IAF: $r_{(24)}$= -.567, p=.003).

A post-hoc power analysis revealed a power of 0.83 for the correlational analysis in the TD group examining the relationship between IAF and the illusory rates. This power is higher as compared with previous studies that investigated the relationship between IAF and AV integration with similar paradigm in neurotypical adults: for example, the post-hoc power in Cecere et al., (2015) was 0.7, in Cooke et al. (2019) was 0.52 and in Venkus and Hughes (2021) was 0.32. We are aware, however, that post-hoc power analyses sometimes can be misleading (Zhang et al, 2019; Gillet et al., 1994) and must be taken cautiously.

[Insert Figure 4 about here]

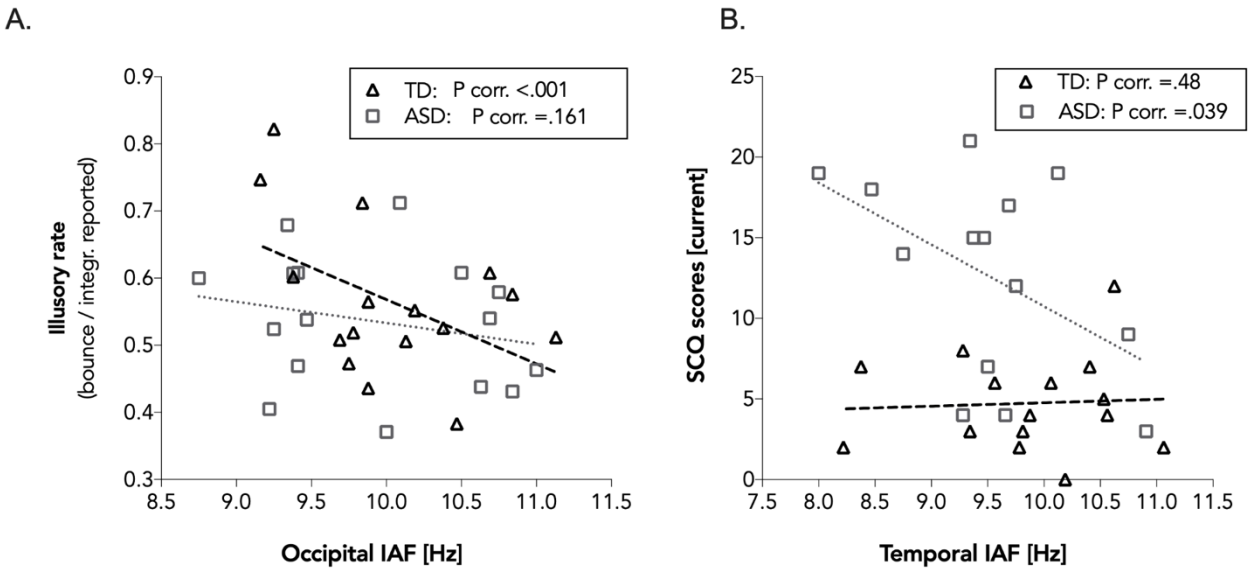


Figure 4. Partial negative correlations of the individual alpha peak (IAF) with the illusory rate and the ASD symptomatology. (A) TD group shows a negative correlation between the illusory rate and the individual alpha peak

in the occipital cluster: the slower the alpha rhythm, the more pronounced the illusory effect, indexing a temporal widening of the AV binding window. (B) Inverse relationship between the ASD symptomatology measured with the SCQ and the alpha peak in the temporal clusters for the group of children with ASD.

Discussion

This study represents, to the best of our knowledge, the first attempt to elucidate the psychophysiological mechanisms behind the altered multisensory processing in ASD within an oscillatory framework. The bistable audiovisual paradigm used here has the advantage of discerning the correlates of AV integration/segregation dynamics, while keeping constant the physical stimulation. Our results showed that pathways driving AV integration/segregation dynamics differentiated children with ASD from their TD peers already at the ERPs level. We observed a significant modulation of peri-stimulus mean amplitude in visual areas for trials where AV stimuli were bounded together (bounce) or processed separately (stream) in the TD group. This modulation was not found in the ASD group, neither in the occipital nor in the temporal areas.

In addition to ERPs, we investigated the neural mechanisms underlying AV integration/segregation in ASD through an extensive characterization of the oscillatory dynamics of both endogenous and task-related neural activity. In TD children, we found a set of results that was largely comparable to the evidence described in the neurotypical adult population (Keil and Senkowski, 2018; Bauer et al., 2020). The frequency of the individual alpha rhythm measured at rest (i.e. the IAF peak) was negatively related to the illusory rate, suggesting that within the TD group a faster alpha rhythm predicted less AV integration (i.e., increased AV segregation) and thus a lower susceptibility to the illusion, in agreement with previous reports using similar paradigms in neurotypical adults (Cecere et al., 2015; Cooke et al., 2019; Venksus and Hughes, 2021). In terms of task-related oscillatory activity, the TD group showed a significant differentiation of cross-modal integration/segregation indexed by theta power, with an increased peri-stimulus amplitude of occipital theta oscillations in bounce (integration) trials as compared to stream (segregation) trials. This result also converges with previous studies in neurotypical adults employing the McGurk illusion to address AV integration

555 (Morís Fernández et al., 2017). Another sign of anticipatory activity predicting AV integration in TD
556 children was the phase of pre-stimulus theta oscillations in the left auditory regions, which was
557 significantly clustered around a common phase angle for trials in which participants reported AV
558 integration (i.e., bounce percept). Overall, our results are consistent with the idea that not only the
559 length of the oscillatory cycle (an endogenous and largely stable trait) but also the phase and
560 amplitude of neural oscillations in the pre- and peri- stimulus period (a variable state) determine the
561 cross-modal dynamics between segregation and integration.
562 Although EEG signatures reflecting AV integration/segregation processes were evident also in the
563 ASD group, their nature was markedly different as compared to the TD group. First, there was no
564 significant correlation between the IAF peak and the illusory rate, despite a similar IAF mean value
565 between groups. Considering that resting-state alpha has been proposed as a key mechanism for
566 temporal sampling within and across sensory modalities (Cecere et al., 2015; Samaha and Postle,
567 2015; Keil and Senkowski, 2017; Ronconi et al., 2018; Ghiani et al., 2021; Cooke et al., 2019;
568 Venksus and Hughes, 2021), the absence of a significant correlation within the ASD group may
569 suggest a less stable - or at least noisier - cortical alpha sampling mechanism. However, it remains
570 the possibility that this null effect was driven by a higher heterogeneity in the ASD group that would
571 require a larger sample size to properly address the relationship between alpha peak and the AV
572 integration rate.
573 Concerning the task-related oscillatory activity, the ASD group contrarily to their TD peers did not
574 show neither signs of pre-stimulus inter-trials phase concentration differences in temporal regions,
575 nor peri-stimulus theta amplitude modulation in occipital regions. Theta oscillations have been
576 proposed to index the formation of prediction of the upcoming auditory stimulus based on visual
577 dynamic information, for example in studies on neurotypical adults investigating the combination of
578 lips movement and perceived phonemes, possibly reflecting communication between visual, auditory,
579 and frontal areas (Keil and Senkowski, 2018). According to our results, it seems that participants with
580 ASD do not rely on anticipatory mechanisms that typically track the dynamic of visual events (in our

581 case the moving objects) to predict the time of their salient changes (in our case the instant when the
582 collision occurs) and to anticipate the onset or change of auditory stimuli.
583 Results of the post-stimulus oscillatory activity seem to converge toward this interpretation and also
584 add further important elements for understanding AV integration anomalies in ASD. In terms of phase
585 reset of neural oscillations, we observed in our data that the timing of the highest phase concentration
586 (ITC) was strictly dependent on the timing of the auditory event. Indeed, the maximum phase
587 coherence in the pre-coincidence condition was anticipated in time relative to the synchronous
588 condition, that in turn was anticipated relative to the post-coincidence condition. This internal control
589 speaks in favor of our interpretation of this effect as truly reflecting an auditory-induced phase reset.
590 Importantly, when looking at differences between groups, we found that the auditory-induced phase
591 reset of oscillations in the visual cortex was a key marker of differentiation between AV integration
592 and segregation in the ASD group. Specifically, a stronger occipital phase reset of oscillations in the
593 alpha band was evident in segregation as compared to integration trials, as if for the ASD group the
594 sound-induced phase alignment in the occipital regions supports a more accurate segregation of
595 audio-visual events, breaking AV integration and the illusory outcome. This result is suggesting an
596 opposite modulation relative to what we know from the neurotypical literature, where auditory-
597 induced phase reset is supposed to be important for opening a temporal window where audio-visual
598 integration can occur (Mercier et al., 2015; Romei et al., 2012), a pattern that emerged also in the
599 present study in TD children in the post-coincidence condition (and, albeit not significant, in the
600 synchronous condition). On the contrary, results observed in the ASD group suggest that a stronger
601 auditory-induced phase reset promotes the parsing of audiovisual stimuli. At a more speculative level,
602 it seems that a disproportionate neural processing of the sound in ASD – which in turn would lead to
603 a stronger cross-modal (auditory-to-visual) influence – can reduce AV integration and promote,
604 instead, segregation.
605 Such speculation seems to find support in the results of beta oscillations in temporal areas in the peri-
606 stimulus period. Indeed, contrary to the TD group, in the ASD group there was an increased power

of high alpha and low beta (12-20 Hz) oscillations in temporal channels during AV integration. Considering the consistent body of evidence showing that the power in this frequency range is inversely related to the cortical activation (i.e. higher alpha/beta power indexes an increased cortical inhibition; e.g. Jensen et al., 2012; Klimesch, 2012 for reviews), a candidate hypothesis is that trials with reduced cortical processing of auditory stimulus were more likely to lead to AV integration.

The two results just described, showing that AV integration in ASD emerges in the presence of a reduced auditory-induced phase reset in occipital areas and a higher alpha/low beta oscillatory amplitude (i.e. an index of cortical inhibition) in temporal areas, are coherently pointing toward the critical role of auditory signals in regulating multisensory dynamics in ASD. Specifically, they suggest that a stronger reactivity to auditory inputs, which has been largely explored in individuals with ASD (for a review see Williams et al., 2021), can be one of the main factor behind AV integration anomalies in ASD.

Lastly, we found an unexpected correlation between ASD social and communication difficulties and the endogenous alpha rhythms in auditory regions. Specifically, individuals with more severe ASD manifestations were characterized by a slower alpha rhythm (i.e. higher IAF peak) in auditory areas, pointing to a potential link between the sampling frequency in the auditory modality and some of the core symptoms of ASD. However, given the unclear link between IAF in temporal regions and the AV integration, this result should be taken cautiously as it would require further investigations.

Altogether, our results fit nicely with the idea that multisensory processing is not a monolithic operation, but rather it could be described as a family of computations that ultimately result in distinct perceptual outcomes. Multisensory processing reflects the highly flexible and adaptive weighting of lower-level (e.g., physical characteristics of the stimuli) and higher-level (e.g., expectation) factors (Van Atteveldt et al., 2014; Murray et al., 2016). The bistable nature of our paradigm allows to reveal that the relative contribution of multiple neural mechanisms to AV integration/segregation dynamics in TD children is largely comparable to the one described in the neurotypical adult literature, suggesting that some of the key neural computations driving cross-modal integration/segregation may

633 be similar throughout the neurotypical development. Notably, the convergence of our
634 neurophysiological results at distinct level of analysis points out the relevance of an endogenous trait
635 measured at rest (the individual frequency of the alpha sampling), and of the anticipatory oscillations
636 at single-trial level for driving the perceptual outcome in neurotypical populations.
637 In contrast, children with ASD seem to weigh differently the distinct computations supporting
638 multisensory processing. The marginal reliance on anticipatory signals of AV co-occurrence was
639 counterbalanced by a greater impact of post-stimulus neural dynamics. Although we cannot infer that
640 the neural mechanisms behind the anomalies described here in AV segregation/integration are
641 common to other multisensory domains (e.g. visuo-tactile) or to unisensory segregation/integration
642 processes, these data suggest that children with ASD rely on *different* neural computations to support
643 integration/segregation of visual and auditory information. This peculiar set of "sampling priorities"
644 suggests intriguing clinical and research implications, considering that a future challenge may be to
645 further clarify the neural oscillatory framework supporting multisensory dynamics in more complex
646 (e.g. socially relevant) stimuli.
647 In conclusions, starting from the assumption that the balance between integration and segregation in
648 multisensory processes may play a crucial role in supporting high-level social functions such as
649 speech comprehension, by using a bistable audiovisual paradigm and EEG, we found that in
650 individuals with ASD, differently from the TD group, audiovisual integration could be predicted
651 neither by the speed of endogenous alpha rhythms nor by the amplitude or phase of pre-stimulus
652 low/mid frequencies oscillations. Instead, a different neural processing of the sound and a differential
653 post-stimulus sound-induced realignment of visual cortical oscillations could explain integration vs.
654 segregation in ASD. These results show how multisensory integration depends on different
655 endogenous, anticipatory and stimulus-driven factors, which seem selectively altered in ASD.
656

Acknowledgments

This work was supported by the "5per1000" funds for biomedical research (Scientific Institute IRCCS Medea; 2016, 2017, 2019 to LC), and by the Italian Ministry of Health (RC 2016-2018 to LC). The funders did not participate in the conception and development of this work.

The contribution of the clinical staff of the Scientific Institute IRCCS Medea as well as of children and their families are gratefully acknowledged.

Financial disclosure section

None of the authors reported financial interests or potential conflicts of interest.

References

American Psychiatric Association (2013) DSM-V.

Ainsworth, K., Ostrolenk, A., Irion, C., & Bertone, A. (2021). Reduced multisensory facilitation exists at different periods of development in autism. Cortex, 134, 195-206.

Bauer AKR, Debener S, Nobre AC (2020) Synchronisation of Neural Oscillations and Cross-modal Influences. Trends Cogn Sci 24:481–495.

Beker S, Foxe JJ, Molholm S (2018) Ripe for solution: Delayed development of multisensory processing in autism and its remediation. Neurosci Biobehav Rev 84:182–192.

Berens P (2009) Journal of Statistical Software CircStat: A MATLAB Toolbox for Circular Statistics.

Brainard DH (1997) The Psychophysics Toolbox. Spat Vis 10:433–436.

Busch NA, Dubois J, VanRullen R (2009) The phase of ongoing EEG oscillations predicts visual perception. J Neurosci 29:7869–7876.

Cecere R, Rees G, Romei V (2015) Individual differences in alpha frequency drive crossmodal illusory perception. Curr Biol 25:231–235.

Chan, J. S., Langer, A., & Kaiser, J. (2016). Temporal integration of multisensory stimuli in autism spectrum disorder: a predictive coding perspective. Journal of Neural Transmission, 123(8), 917-923.

Collignon, O., Charbonneau, G., Peters, F., Nassim, M., Lassonde, M., Lepore, F., ... Bertone, A. (2013). Reduced multisensory facilitation in persons with autism. Cortex, 49(6), 1704–1710.

Cooke J, Poch C, Gillmeister H, Costantini M, Romei V. (2019) Oscillatory Properties of Functional Connections Between Sensory Areas Mediate Cross-Modal Illusory Perception. J Neurosci.39(29): 5711-5718.

Delorme A, Makeig S (2004) EEGLAB: an open sorce toolbox for analysis of single-trail EEG dynamics including independent component anlaysis. J Neurosci Methods 134:9–21.

Feldman JI, Dunham K, Cassidy M, Wallace MT, Liu Y, Woynaroski TG (2018) Audiovisual

multisensory integration in individuals with autism spectrum disorder: A systematic review and meta-analysis. Neurosci Biobehav Rev 95:220–234.

Ghiani A, Maniglia M, Battaglini L, Melcher D, Ronconi L (2021) Binding Mechanisms in Visual Perception and Their Link With Neural Oscillations: A Review of Evidence From tACS. Front Psychol 12.

Gillett, R. (1994). Post hoc power analysis. Journal of Applied Psychology, 79: 783.

Han B, VanRullen R (2017) The rhythms of predictive coding? Pre-stimulus phase modulates the influence of shape perception on luminance judgments. Sci Rep 7:1–10.

Jensen O, Bonnefond M, VanRullen R (2012) An oscillatory mechanism for prioritizing salient unattended stimuli. Trends Cogn Sci 16:200–206.

Kambe, J., Kakimoto, Y., & Araki, O. (2015). Phase reset affects auditory-visual simultaneity judgment. Cognitive neurodynamics, 9(5), 487-493.

Keil J, Senkowski D (2017) Individual alpha frequency relates to the sound-induced flash illusion. Multisens Res 30:565–578.

Keil J, Senkowski D (2018) Neural Oscillations Orchestrate Multisensory Processing. Neurosci:107385841875535.

Klimesch W (2012) Alpha-band oscillations, attention, and controlled access to stored information. Trends Cogn Sci 16:606–617.

Lai MC, Anagnostou E, Wiznitzer M, Allison C, Baron-Cohen S (2020) Evidence-based support for autistic people across the lifespan: maximising potential, minimising barriers, and optimising the person–environment fit. Lancet Neurol 19:434–451.

Lakatos P, O'Connell MN, Barczak A, Mills A, Javitt DC, Schroeder CE (2009) The Leading Sense: Supramodal Control of Neurophysiological Context by Attention. Neuron 64:419–430.

Lopez-Calderon J, Luck SJ (2014) ERPLAB: An open-source toolbox for the analysis of event-related potentials. Front Hum Neurosci 8:213.

Lord C, Rutter M, DiLavore PC, Risi S (2002) Autism diagnostic observation schedule: ADOS.

718 Western Psychological Services.
719 Makeig, S., Westerfield, M., Jung, T. P., Enghoff, S., Townsend, J., Courchesne, E., & Sejnowski,
720 T. J. (2002). Dynamic brain sources of visual evoked responses. Science, 295(5555), 690-694.
721 Maris E, Oostenveld R (2007) Nonparametric statistical testing of EEG- and MEG-data. J Neurosci
722 Methods 164:177–190.
723 Mercier MR, Foxe JJ, Fiebelkorn IC, Butler JS, Schwartz TH, Molholm S (2013) Auditory-driven
724 phase reset in visual cortex: Human electrocorticography reveals mechanisms of early
725 multisensory integration. Neuroimage 79:19–29.
726 Mercier MR, Molholm S, Fiebelkorn IC, Butler JS, Schwartz TH, Foxe JJ (2015) Neuro-oscillatory
727 phase alignment drives speeded multisensory response times: An electro-corticographic
728 investigation. J Neurosci 35:8546–8557.
729 Morillon B, Schroeder CE (2015) Neuronal oscillations as a mechanistic substrate of auditory
730 temporal prediction. Ann N Y Acad Sci 1337:26–31.
731 Morís Fernández L, Macaluso E, Soto-Faraco S (2017) Audiovisual integration as conflict
732 resolution: The conflict of the McGurk illusion. Hum Brain Mapp 38:5691–5705.
733 Murray MM, Lewkowicz DJ, Amedi A, Wallace MT (2016) Multisensory Processes: A Balancing
734 Act across the Lifespan. Trends Neurosci 39:567–579.
735 Pelli DG (1997) The VideoToolbox software for visual psychophysics: Transforming numbers into
736 movies. Spat Vis 10:437–442.
737 Pöppel E (2009) Pre-semantically defined temporal windows for cognitive processing. Philos Trans
738 R Soc B Biol Sci 364:1887–1896.
739 Robertson CE, Baron-Cohen S (2017) Sensory perception in autism. Nat Rev Neurosci 18:671–684
740 Available at: http://dx.doi.org/10.1038/nrn.2017.112.
741 Romei V, Gross J, Thut G (2012) Sounds reset rhythms of visual cortex and corresponding human
742 visual perception. Curr Biol 22:807–813.
743 Ronconi L, Busch NA, Melcher D (2018) Alpha-band sensory entrainment alters the duration of

temporal windows in visual perception. Sci Rep 8:1–10.

Samaha J, Postle BR (2015) The Speed of Alpha-Band Oscillations Predicts the Temporal Resolution of Visual Perception. Curr Biol 25:2985–2990.

Sekuler R, Sekuler AB, Lau R (1997) Sound alters visual motion perception. Nature 385:308.

Spence C, Squire S (2003) Multisensory integration: Maintaining the perception of synchrony. Curr Biol 13.

Stevenson, R. A., Siemann, J. K., Schneider, B. C., Eberly, H. E., Woynaroski, T. G., Camarata, S. M., & Wallace, M. T. (2014). Multisensory temporal integration in autism spectrum disorders. *The Journal of Neuroscience*, *34*(3), 691–697.

Stevenson, R. A., Segers, M., Ferber, S., Barense, M. D., Camarata, S., & Wallace, M. T. (2016). Keeping time in the brain: Autism spectrum disorder and audiovisual temporal processing. Autism Research, 9(7), 720–738. https://doi.org/10.1002/aur.1566.

Tallon-Baudry C, Bertrand O, Delpuech C, Pernier J (1996) Stimulus specificity of phase-locked and non-phase-locked 40 Hz visual responses in human. J Neurosci 16:4240–4249.

Ten Oever S, Sack AT (2015) Oscillatory phase shapes syllable perception. Proc Natl Acad Sci U S A 112:15833–15837.

Thézé R, Giraud AL, Mégevand P (2020) The phase of cortical oscillations determines the perceptual fate of visual cues in naturalistic audiovisual speech.

Van Atteveldt N, Murray MM, Thut G, Schroeder CE (2014) Multisensory integration: Flexible use of general operations. Neuron 81:1240–1253.

VanRullen R (2016a) Perceptual Cycles. Trends Cogn Sci 20:723–735.

VanRullen, R. (2016b). How to evaluate phase differences between trial groups in ongoing electrophysiological signals. *Frontiers in neuroscience*, *10*, 426.

Venskus A, Hughes G. (2021) Individual differences in alpha frequency are associated with the time window of multisensory integration, but not time perception. Neuropsychologia. 20; 159:107919. doi: 10.1016/j.neuropsychologia.2021.107919.

Wallace MT, Stevenson RA (2014) The construct of the multisensory temporal binding window and its dysregulation in developmental disabilities. Neuropsychologia 64:105–123.

Wechsler D (1993) WISC-R. Scala d'intelligenza Wechsler per bambini riveduta.

Williams, Z. J., He, J. L., Cascio, C. J., & Woynaroski, T. G. (2021). A review of decreased sound tolerance in autism: Definitions, phenomenology, and potential mechanisms. *Neuroscience & Biobehavioral Reviews*. https://doi.org/10.1016/j.neubiorev.2020.11.030

Zhang, Y., Hedo, R., Rivera, A., Rull, R., Richardson, S., & Tu, X. M. (2019) Post hoc power analysis: is it an informative and meaningful analysis? General psychiatry, 32: e100069.

Zhou H yu, Cai X lu, Weigl M, Bang P, Cheung EFC, Chan RCK (2018) Multisensory temporal binding window in autism spectrum disorders and schizophrenia spectrum disorders: A systematic review and meta-analysis. Neurosci Biobehav Rev 86:66–76.